

### Exploring the use of web searches for risk communication during COVID-19 in Germany

Kristensen, Kaja; Lorenz, Eva; May, Jürgen; Strauss, Ricardo

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

#### Empfohlene Zitierung / Suggested Citation:

Kristensen, K., Lorenz, E., May, J., & Strauss, R. (2021). Exploring the use of web searches for risk communication during COVID-19 in Germany. *Scientific Reports*, 11, 1-10. <https://doi.org/10.1038/s41598-021-85873-4>

#### Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:

<https://creativecommons.org/licenses/by/4.0/deed.de>

#### Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see:

<https://creativecommons.org/licenses/by/4.0>



OPEN

## Exploring the use of web searches for risk communication during COVID-19 in Germany

Kaja Kristensen<sup>1</sup>, Eva Lorenz<sup>2</sup>, Jürgen May<sup>1</sup> & Ricardo Strauss<sup>2</sup>

Risk communication during pandemics is an element of utmost importance. Understanding the level of public attention—a prerequisite for effective communication—implicates expensive and time-consuming surveys. We hypothesise that the relative search volume from Google Trends could be used as an indicator of public attention of a disease and its prevention measures. The search terms 'RKI' (Robert Koch Institute, national public health authority in Germany), 'corona' and 'protective mask' in German language were shortlisted. Cross-correlations between these terms and the reported cases from 15 February to 27 April were conducted for each German federal state. The findings were contrasted against a timeline of official communications concerning COVID-19. The highest correlations of the term 'RKI' with reported COVID-19 cases were found between lags of – 2 and – 12 days, meaning web searches were already performed from 2 to 12 days before case numbers increased. A similar pattern was seen for the term 'corona'. Cross-correlations indicated that most searches on 'protective mask' were performed from 6 to 12 days after the peak of cases. The results for the term 'protective mask' indicate a degree of confusion in the population. This is supported by conflicting recommendations to wear face masks during the first wave. The relative search volumes could be a useful tool to provide timely and location-specific information on public attention for risk communication.

COVID-19 (Corona Virus Disease 2019) is caused by the Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV-2). The global spread of the virus led to COVID-19 being classified as a pandemic in March 2020, affecting the lives of billions of people. At the time of writing, over 29 million confirmed cases and nearly 950,000 deaths have been reported worldwide<sup>1</sup>. In Germany, ca. 260,000 confirmed cases and 9400 deaths have been reported during the first wave<sup>2</sup>.

A distinctive feature of infectious diseases is that individual behaviour can also impact the health of others. Consequently, human behaviour plays a major role in the research and control of infectious diseases<sup>3–5</sup>. In the case of COVID-19, preventive measures to contain the spread of the virus and reduce the burden on health systems are highly relevant given its high secondary attack rate and outstanding deadly toll. In addition to government-imposed contact restrictions and other social distancing policies, a number of individual hygiene measures, i.e. regular hand washing and wearing masks, are recommended in order to contain the spread and to protect oneself<sup>6</sup>.

The risks and uncertainties of emerging infectious diseases can trigger a diverse array of emotional, cognitive, and behavioural responses and affect public behaviour in both constructive (e.g. adopting individual hygiene measures) and destructive manners (e.g. excessive use of health care services)<sup>7–9</sup>. Risk communication, defined as “the exchange of information among interested parties about the nature, magnitude, significance, or control of a risk”<sup>10</sup>, aims at increasing the capacity of the public to act as an effective response partner by encouraging constructive responses amongst the population, i.e. by adopting desired prevention measures<sup>11,12</sup>. Behaviour change can be induced through risk communication by presenting a threat and describing a behaviour change that can mitigate the threat<sup>13</sup>. For supporting effective risk communication it is important to understand how critical health risk information is disseminated and how the public accesses, processes and responds to this information<sup>14</sup>. Research on risk communication has already provided numerous insights thereto. For example, messages are more effective when delivered clear and simple, appeal to reason and emotion, and are tailored to the needs, values, cultural background and experiences of the target audience<sup>13</sup>. In addition, the characteristics

<sup>1</sup>Department of Medical Psychology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. <sup>2</sup>Department of Infectious Disease Epidemiology, Bernhard Nocht Institute for Tropical Medicine, Hamburg, Germany. ✉email: ricardo.strauss@bnitm.de

of the entity sending a message are likewise important, for example, credibility of authority, transparency and perceived expertise influence the persuasiveness of the communication<sup>15</sup>.

Important prerequisites for effective communication, i.e. communication that leads to the desired constructive response, is described in the Protective Action Decision Model. According to this Model, the target population must (a) be exposed to the information, (b) pay attention to it and (c) understand it<sup>16</sup>. Measuring the degree of exposure, attention and comprehension of communications would usually require time-consuming and resource-intensive data collection methods. Particularly in the case of public attention, conventional data collection methods such as surveys are criticised for not being able to adequately capture the dynamic nature of public attention<sup>17</sup>. The vast amount of real-time information contained on the internet could open up innovative ways to assess public attention<sup>17</sup>. Web searches are not only a valuable resource for individuals seeking information, but also for the scientific community, since search queries contain geospatial and temporal data.

The Internet represents a common source for finding health-related information<sup>18</sup>. Numerous studies have shown how analysing web search queries can assist in describing and predicting outbreaks especially in the context of sparse data<sup>19–21</sup>. This approach has also been used for COVID-19 already<sup>22,23</sup>. However, web searches have also been analysed to describe public attention in relation to a public health issue. In the area of non-communicable diseases, studies can be found that examine public attention to preventive measures such as cancer screening or smoking cessation<sup>24–27</sup>. In the area of infectious diseases, studies mainly examine public attention to the disease itself<sup>9,28–31</sup>, but further examine attention to specific prevention measures<sup>32–35</sup>. Public interest in key institutions during a pandemic providing information on who is involved in risk communication, namely public health authorities, has not yet been studied.

The aim of this study is to combine these approaches and to investigate the usefulness of clustered web searches as an indicator of public attention to the disease itself and to prevention-relevant aspects in order to inform risk communication strategies in the context of the ongoing pandemic in Germany.

## Materials and methods

**Data sources.** *Officially reported cases.* Daily incidence data of officially reported COVID-19 cases in Germany were obtained from the Robert Koch Institute (RKI) which is the national public health authority in Germany. Data for the first wave was used, which corresponds to the period from 15 February 2020 to 27 April 2020. Data are publicly available ([https://npgeo-corona-npgeo-de.hub.arcgis.com/datasets/ef4b445a53c1406892257fe63129a8ea\\_0?geometry=-19.624%2C46.270%2C35.879%2C55.886](https://npgeo-corona-npgeo-de.hub.arcgis.com/datasets/ef4b445a53c1406892257fe63129a8ea_0?geometry=-19.624%2C46.270%2C35.879%2C55.886)). The daily number of new cases were retrieved for each of the 16 federal states in Germany separately as well as clustered national data.

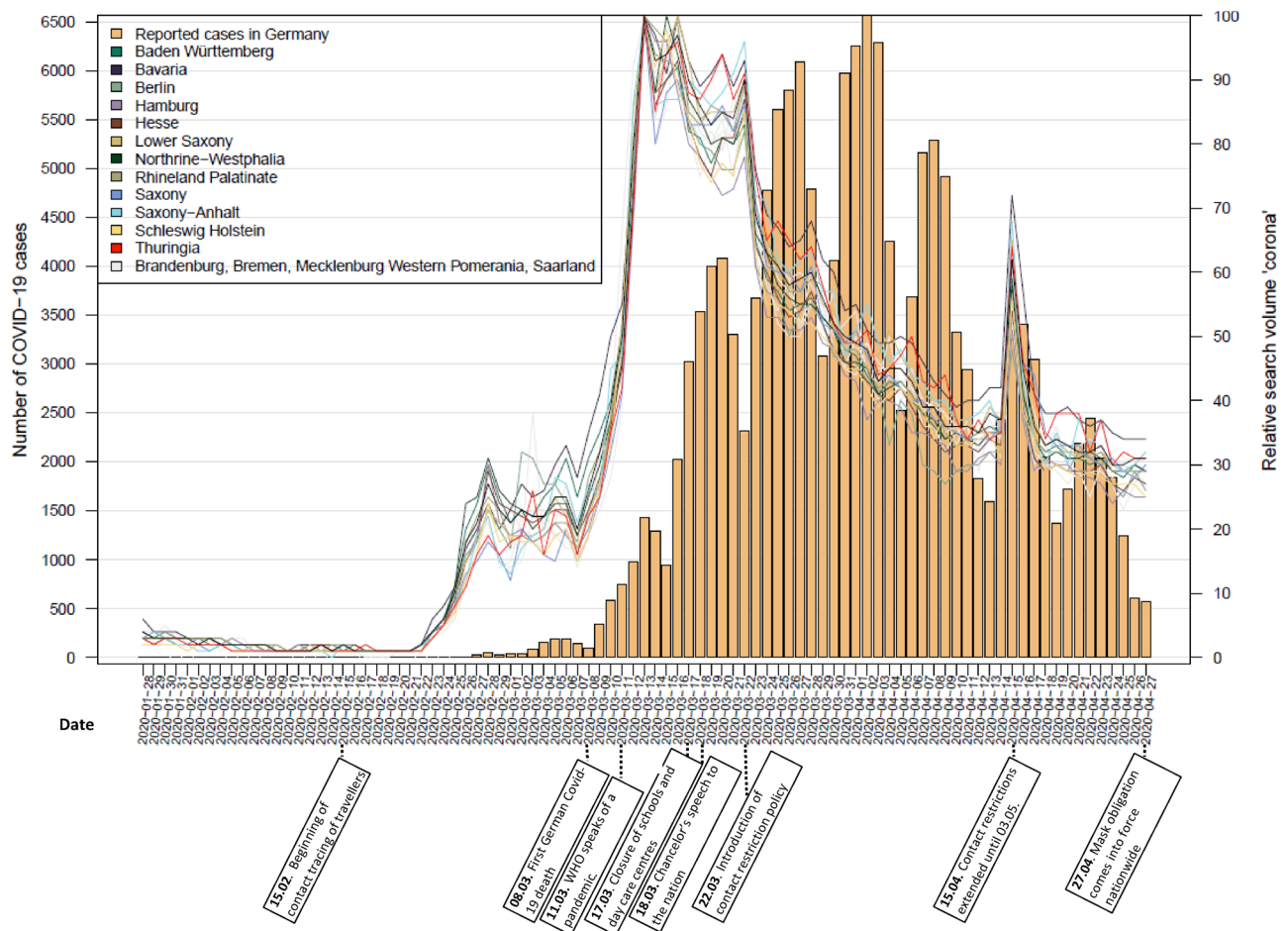
*Web search queries.* Google Trends provides access to Internet search patterns by analysing web queries on the Google search browser and other affiliated Google sites. It provides a Relative Search Volume (RSV), i.e. the query share of a particular term for a given location and time, normalised by the highest query share of that term over the given time period. Data is presented on a relative scale from 0 to 100, with 0 representing no related queries and 100 representing the maximum number of related queries.

During the period of 15 February to 27 April, various German query terms used for seeking information on COVID-19 and its prevention measures were investigated. Search terms were collected and divided into the following thematically related groups: the virus or the disease itself, sources of information and preventive measures, which included hand hygiene and facial masks. After this, search terms of each group were jointly queried on Google Trends to identify the most frequently searched terms within a thematic group. Search terms related to the disease or virus, respectively were ‘corona’, ‘coronavirus’, ‘covid-19’ and ‘covid 19’. Search queries related to information sources were ‘corona podcast’, ‘corona info’, ‘corona hotline’, ‘Robert Koch Institut’, ‘RKI’, and ‘Gesundheitsamt’, whilst search queries for hand hygiene and masks were ‘Händewaschen’, ‘Hände waschen’, ‘Händedesinfektion’, ‘Händedesinfizieren’ and ‘Hände desinfizieren’, and ‘Mundschutz’, ‘Maske’, ‘Gesichtsmaske’, ‘Schutzmaske’, ‘Stoffmaske’, respectively. The most frequently searched terms were then selected. Finally, daily data from 15 February to 27 April 2020 on web searches for ‘RKI’, ‘Mundschutz’ (protective mask) and ‘corona’ were obtained separately from Google Trends (<https://trends.google.com/trends/>) on 01 May 2020. Data were retrieved for each federal state in Germany and nationally clustered for the whole country. Thereafter, data was downloaded in csv-format. Search queries related to hand hygiene were not suitable for further statistical analyses because the RSV only displayed extreme values around 0 or 100 with little variation in between. Therefore, this group of search queries was excluded from subsequent analyses.

*Risk communication milestones.* Important announcements in the period from 15 February to 27 April were retrieved from timelines published by the German Ministry of Health (MoH) (<https://www.bundesgesundheitsministerium.de/coronavirus/chronik-coronavirus.html>) and the World Health Organization (WHO) (<https://www.who.int/news-room/detail/29-06-2020-covidtimeline>). In addition, short reports during the same period from a major national public media network (<https://www.tagesschau.de/faktenfinder/hintergrund/corona-chronik-pandemie-101.html>) were screened.

**Data analysis.** Statistical analyses were conducted using SPSS version 25 and R version 3.6.2. Line plots in combination with histograms, and cross-correlations were generated to assess the association between Google search trends and the officially reported COVID-19 cases.

To characterise time dependence between time-series from two different data sources, cross-correlations were used. The time dependence between two time-series is termed as lag and indicates the direction of the two time-series correlated. A lag of – 1 suggests that a peak in Google Trends precedes a peak in the officially reported COVID-19 cases by 1 day, and vice versa for positive lags.



**Figure 1.** Nationally reported COVID-19 cases and RSV for the search term 'corona' at state level.

Both time series were normalised to account for the different units and scales (RSV and absolute number of daily new cases). Thresholds for interpreting a correlation coefficient suggested by Hinkle et al. were applied<sup>36</sup>.

The risk communication milestones were reviewed based on their content and relevance for the present study by two independent reviewers. Disagreement was solved by discussion.

**Ethics declarations.** Ethical approval and consent to participate were not necessary as the study was based on openly available aggregated data.

## Results

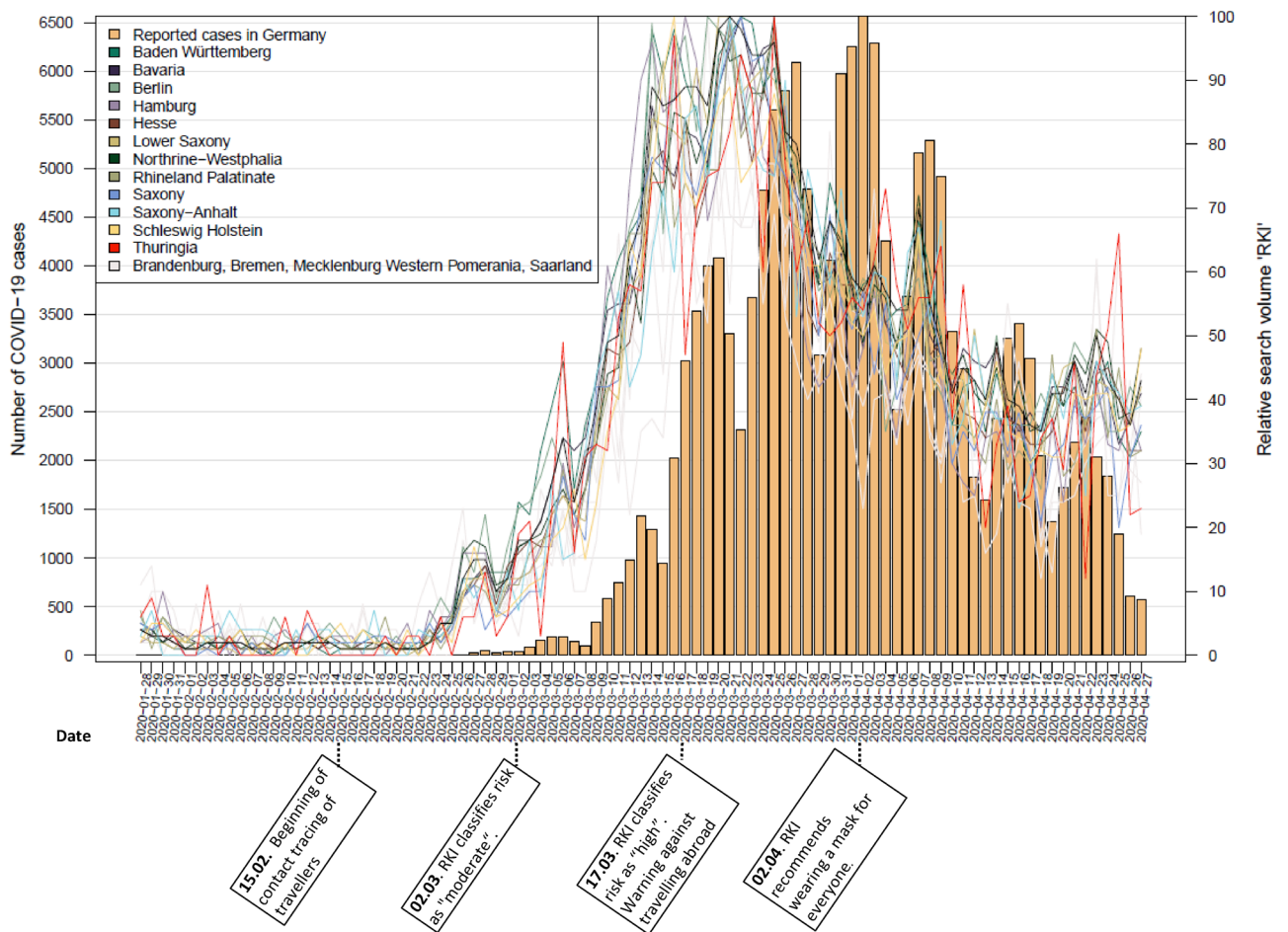
**Time trends.** Histograms of officially reported cases on a national level in combination with line graphs of Google search queries for 'RKI', 'corona' and 'protective masks' at state level are shown in Figs. 1, 2 and 3. Most relevant communications and announcements are also presented in these figures. Figures 1 and 2 show that in all German federal states the RSV of the search terms 'RKI' and 'corona' reach their peak (mid-March) before the number of cases peak (beginning of April). The RSV of 'corona' shows a second peak of searches in mid-April. Meanwhile the RSV of 'protective mask' presents with three peaks with the highest peak at a point where the curve of confirmed cases is already flattening (end of April) (Fig. 3).

**Time-series cross-correlations.** The linear association between officially reported COVID-19 cases and Google Trends patterns for the search terms 'RKI', 'corona' and 'protective' mask were assessed using cross-correlations, as shown in Figs. 4, 5 and 6.

Visual assessment of Figs. 4 and 5, which refer to the search terms 'corona' and 'RKI' respectively, reveals that the correlations are high in most federal states with the exception of Brandenburg, Bremen, Mecklenburg-Western Pomerania and Saarland. These federal states are described separately from the rest in the following.

For the group of federal states with high correlation, the RSV of the term 'corona' correlates highly with officially reported cases ( $r = 0.707$  to  $r = 0.804$ ) at lags between  $-5$  and  $-12$  days. In the same group of states, the RSV for 'RKI' and reported cases correlate highly ( $r = 0.711$  to  $r = 0.856$ ) when searches precede the reported cases between 2 and 10 days (range of lags:  $-2$  to  $-10$  days).

In the federal states Brandenburg, Bremen, Mecklenburg-Western Pomerania and Saarland only moderate correlations ( $r = 0.453$  to  $r = 0.688$ ) are seen between the RSV of the term 'corona' and reported cases. Correlations



**Figure 2.** Nationally reported COVID-19 cases and RSV for the search term 'RKI' at state level.

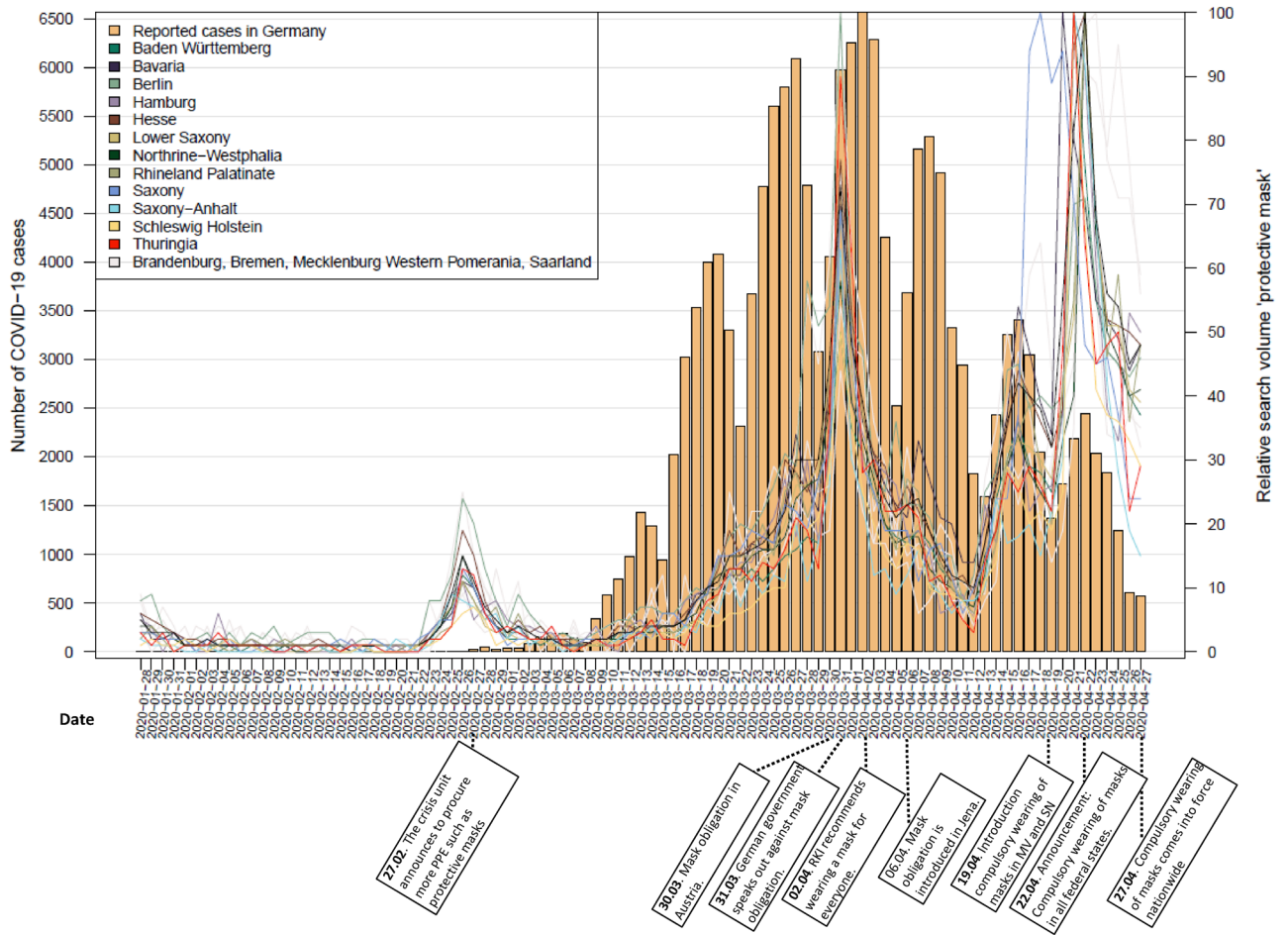
are highest when the searches precede the reported cases (range of lags:  $-5$  to  $14$  days). For these states, a similar trend can be observed for the term 'RKI' ( $r = 0.449$  and  $r = 0.682$ ; range of lags:  $-10$  to  $0$  days).

A different picture results from the search term 'protective mask' (Fig. 6). In contrast to the other two search terms, the highest correlation coefficients ( $r = 0.400$  to  $r = 0.602$ ) result when reported cases precede the RSV for 'protective mask' by 5 to 14 days (range of lag  $+5$  to  $+14$  days).

**Risk communication milestones and time trends of search queries.** Following the timeline of the data period (15 February to 27 April) chronologically (Fig. 7), an important announcement was made on 02 March when the RKI classified the risk as "moderate". Just over a week later, on 11 March, COVID-19 was classified as a pandemic by WHO. During this period, searches for 'RKI' and 'corona' began to soar (Figs. 1, 2). On 13 March, most of the German states announced that they would temporarily close the schools. On 17 March, the RKI classified the risk as "high" and all schools and day care centres were closed. At that time, the search queries for the terms 'corona' and 'RKI' both reached their highest level (Figs. 1, 2). Five days later, on 22 March, the Federal Government and the federal states agreed on guidelines for restricting social contacts. Central elements were the prohibition of staying in public places with more than one other person not living in the same household, and the closure of restaurants, bars, and service providers with physical contact, such as hairdressers or massage studios. At this point, there was a renewed increase in search queries of the term 'corona'. The RSV for 'corona' previously showed a decreasing trend.

On 31 March, the German government expressed its opposition to the obligation to wear masks as medical masks should be reserved for health care personnel. At this point, the search queries for the term 'protective mask' increased steeply in all federal states (Fig. 3). On 02 April, the RKI extended its recommendation regarding wearing masks. While wearing masks was previously recommended for people with respiratory diseases, they now recommended wearing masks in public places for the general population. A few days later (06 April), compulsory wearing of masks was introduced in Jena, a city in the federal state of Thuringia. In the following weeks, several federal states introduced compulsory wearing of masks, while there was no nationwide regulation. Meanwhile, it was announced on 15 April that the contact restrictions, announced on 22 March, would be extended until 03 May and that the schools would reopen on 04 May. The previously decreasing trend seen in the RSV of the term 'corona' showed a sudden increase at this point (Fig. 1). On 17 April, the Federal Minister of





**Figure 3.** Nationally reported COVID-19 cases and RSV for the search term ‘protective mask’ at state level.

Health stated on national television that the first wave of infection had been well overcome and that the country should gradually return to normality while staying well-prepared and alert. On 22 April, a nationwide mask obligation was announced which came into force on 27 April. At this point, queries for ‘protective mask’ then reached their peak in the period under review (Fig. 3).

**Discussion**

The findings show that different search terms related to the pandemic had different trends over the course of the first wave. While searches for the terms ‘RKI’ and ‘corona’ reached their peak before the number of cases peaked, the RSV of the term ‘protective mask’ peaked when the curve was already flattening.

The RSV for the term ‘RKI’ followed a similar trend as the RSV for the term ‘corona’, with relatively few searches until the end of February and most searches performed in mid-March with a subsequent decline in searches. At times of increasing attention to ‘corona’, attention to the national health authority ‘RKI’ also increased. This could indicate that the Institute is a trusted source of information for the public. This hypothesis is supported by data from the GESIS Panel Special Survey on the Coronavirus SARS-CoV-2 Outbreak in Germany. The data displays, that the RKI enjoys the greatest public trust in dealing with the pandemic compared to other institutions<sup>37</sup>.

The temporal relationship between the RSV of ‘corona’ and the reported cases suggests that the term has a certain predictive value that could be used in the context of sparse data. A study by Effenberger et al. examined this temporal relationship in different countries and also found that the peaks of web searches for the term ‘coronavirus’ preceded the peaks of the reported cases by 11.5 days on average<sup>38</sup>. However, visual examination of the data from this study also suggests that RSV is closely linked with risk communication milestones, such as categorisation of risk levels, public debates on appropriateness of measures, and policy announcements. The intention to include these milestones alongside the RSV of search queries and official case data is based on the potential influence of risk communication milestones on the way people respond and search for online information. Therefore, risk communication milestones may be considered important and meaningful in the visual representation along with the other two data sources.

The observation related to an increase in web searches on ‘RKI’ and ‘corona’ preceding the increase of reported cases was also described in a similar study that monitored search terms related to COVID-19 such as ‘corona’,

Lag	-14	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0
Baden-Württemberg	0,685	0,736	0,771	0,779	0,761	0,742	0,731	0,73	0,728	0,718	0,689	0,644	0,597	0,544	0,497
Bavaria	0,734	0,772	0,791	0,782	0,763	0,75	0,736	0,72	0,703	0,678	0,631	0,58	0,534	0,489	0,446
Berlin	0,627	0,689	0,734	0,742	0,736	0,73	0,73	0,734	0,746	0,718	0,679	0,626	0,57	0,515	0,462
Brandenburg	0,59	0,613	0,645	0,62	0,599	0,61	0,577	0,565	0,58	0,588	0,571	0,538	0,513	0,469	0,422
Bremen	0,262	0,259	0,315	0,311	0,341	0,357	0,412	0,442	0,452	0,391	0,391	0,375	0,342	0,339	0,293
Hamburg	0,5	0,559	0,601	0,641	0,677	0,704	0,72	0,76	0,804	0,802	0,792	0,762	0,737	0,681	0,627
Hesse	0,597	0,644	0,675	0,701	0,707	0,661	0,665	0,686	0,695	0,699	0,678	0,643	0,593	0,549	0,486
Lower Saxony	0,649	0,699	0,728	0,737	0,726	0,72	0,729	0,733	0,747	0,738	0,711	0,674	0,627	0,581	0,534
Mecklenburg Western Pomerania	0,416	0,457	0,515	0,532	0,515	0,501	0,486	0,53	0,606	0,688	0,678	0,66	0,589	0,531	0,504
Northrhine-Westphalia	0,584	0,623	0,652	0,671	0,677	0,671	0,688	0,711	0,734	0,75	0,747	0,713	0,673	0,64	0,584
Rhineland Palatinate	0,609	0,66	0,703	0,735	0,75	0,749	0,754	0,77	0,773	0,784	0,774	0,751	0,722	0,671	0,617
Saarland	0,687	0,673	0,655	0,66	0,632	0,601	0,562	0,549	0,51	0,479	0,447	0,408	0,373	0,346	0,31
Saxony	0,621	0,676	0,714	0,725	0,709	0,706	0,731	0,749	0,768	0,756	0,715	0,674	0,629	0,597	0,551
Saxony-Anhalt	0,59	0,634	0,666	0,697	0,713	0,705	0,737	0,734	0,771	0,74	0,713	0,683	0,662	0,593	0,536
Schleswig Holstein	0,618	0,668	0,719	0,735	0,724	0,714	0,712	0,698	0,698	0,683	0,647	0,606	0,574	0,531	0,478
Thuringia	0,565	0,608	0,656	0,675	0,663	0,67	0,701	0,724	0,735	0,708	0,661	0,624	0,573	0,557	0,501
Germany	0,672	0,719	0,748	0,757	0,748	0,739	0,735	0,741	0,745	0,735	0,702	0,661	0,615	0,569	0,52

.90 to 1.00	Very high correlation
.70 to .90	High correlation
.50 to .70	Moderate correlation
.30 to .50	Low correlation
.00 to .30	negligible correlation

**Figure 4.** Cross-correlation coefficients displaying the linear association between the RSV of the term ‘corona’ and officially reported COVID-19 cases.

‘handwashing’ and ‘masks’. In contrast to this study, however, web searches for the terms ‘face mask’ and ‘surgical mask’ behaved in the same manner and were performed before most cases were reported<sup>33</sup>.

This leads to the question of why the search term ‘protective mask’ displayed a different trend in Germany. The long absence of a nationwide uniform obligation to wear a mask, while such obligations have already been introduced at a local level, could have led to confusion and an increased the need for information retrieval on this topic on the Internet. In addition, there have been conflicting recommendations from various institutions as to whether masks should be worn by the general public or primarily by healthcare professionals. This public debate and uncertainty are reflected in the RSV of the term ‘protective masks’.

The previous example attests how the RSV of search terms can help to assess the response to announcements and messages disseminated to the target audience. We were able to show that at a time when public attention for ‘corona’ was at its highest, relatively few people informed themselves about prevention measures such as wearing protective masks on the internet. Only when a nationwide obligation to wear masks in public areas was announced, did search queries for the term ‘protective mask’ reach their peak.

In retrospective, this is important information for the risk communication strategy. Clarity in communication, especially on the topic of masks, could have stimulated interest in masks at an earlier stage. It would have been desirable for the public to be made aware of this prevention measure at the peak of the first wave, or preferably earlier, so that it could then obtain information about this preventive measure on the Internet.

The results of this study are subject to some limitations that should be considered. First, the intentions behind the analysed web searches are unclear. Different motives such as seeking health information, updating oneself on current policies, or purchasing face masks might have led individuals to search for the terms and do not necessarily represent certain attitudes towards any preventive measures or the disease itself. Furthermore, it is known based on literature that disease-related web searches are subject to media-driven interest known as “celebrity phenomena”<sup>39</sup>, which refers to increased internet queries for certain terms that are temporarily popular in the media.

Lag	-14	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0
Baden-Württemberg	0,651	0,691	0,743	0,784	0,824	0,830	0,829	0,801	0,785	0,777	0,793	0,776	0,747	0,700	0,653
Bavaria	0,706	0,744	0,775	0,812	0,832	0,827	0,820	0,812	0,782	0,753	0,740	0,720	0,676	0,642	0,602
Berlin	0,643	0,651	0,686	0,707	0,738	0,783	0,800	0,824	0,806	0,768	0,747	0,729	0,710	0,702	0,675
Brandenburg	0,633	0,659	0,626	0,658	0,676	0,632	0,631	0,668	0,643	0,629	0,617	0,587	0,538	0,544	0,528
Bremen	0,276	0,328	0,244	0,337	0,317	0,381	0,347	0,369	0,380	0,362	0,377	0,387	0,422	0,393	0,371
Hamburg	0,528	0,563	0,612	0,657	0,711	0,738	0,779	0,811	0,796	0,802	0,791	0,788	0,787	0,768	0,718
Hesse	0,600	0,641	0,662	0,694	0,732	0,751	0,778	0,780	0,780	0,758	0,754	0,751	0,743	0,718	0,688
Lower Saxony	0,619	0,670	0,697	0,732	0,764	0,789	0,804	0,817	0,809	0,803	0,805	0,789	0,765	0,739	0,717
Mecklenburg Western Pomerania	0,249	0,289	0,336	0,340	0,398	0,447	0,609	0,563	0,535	0,576	0,580	0,542	0,558	0,673	0,647
Northrhine-Westphalia	0,547	0,595	0,628	0,671	0,706	0,753	0,774	0,799	0,789	0,798	0,797	0,813	0,818	0,806	0,781
Rhineland Palatinate	0,546	0,610	0,661	0,704	0,747	0,779	0,803	0,825	0,845	0,843	0,829	0,856	0,832	0,795	0,759
Saarland	0,661	0,679	0,682	0,665	0,682	0,638	0,599	0,550	0,535	0,496	0,435	0,433	0,412	0,374	0,313
Saxony	0,592	0,637	0,713	0,759	0,778	0,795	0,811	0,802	0,785	0,786	0,795	0,767	0,727	0,695	0,648
Saxony-Anhalt	0,488	0,550	0,567	0,607	0,710	0,742	0,755	0,726	0,752	0,789	0,744	0,740	0,795	0,719	0,668
Schleswig Holstein	0,564	0,640	0,670	0,715	0,729	0,761	0,791	0,798	0,797	0,749	0,757	0,733	0,741	0,694	0,672
Thuringia	0,563	0,627	0,690	0,662	0,652	0,666	0,671	0,659	0,684	0,711	0,666	0,645	0,610	0,584	0,598
Germany	0,651	0,697	0,734	0,772	0,803	0,821	0,831	0,834	0,824	0,812	0,805	0,792	0,768	0,740	0,702

.90 to 1.00	Very high correlation
.70 to .90	High correlation
.50 to .70	Moderate correlation
.30 to .50	Low correlation
.00 to .30	negligible correlation

**Figure 5.** Cross-correlation coefficients displaying the linear association between the RSV of the term ‘RKI’ and officially reported COVID-19 cases.

Another important limitation arises from the limited applicability of the analysis method in federal states with low absolute populations, as it was the case in Brandenburg, Bremen, Mecklenburg-Western Pomerania and Saarland. The results showed that correlations between the RSVs and reported cases were only from low to moderate in these states. It is therefore questionable to what extent the presented method can be applied in areas with a low population.

Some search terms that are highly relevant for disease prevention and personal protection, namely search queries related to hand hygiene, could not be included in the analysis due to their data distribution. It is assumed that a reduced search frequency resulted in a distribution of data showing only very low or very high values. Different reasons might explain why these terms were not searched for with more frequency, e.g. no need in the population to inform themselves about these topics. As a final limitation, it should be mentioned that RSVs do not provide information about the absolute number of web searches. Therefore, it is not possible to draw conclusions about the actual number of searches performed, but only about their trends over time.

Nevertheless, in this study we could present a novel tool for the fast assessment of public attention in times of an ongoing pandemic and attest how web searches can also inform risk communication strategies about the attention paid to a risk and different aspects of it.

### Conclusion

In the context of an ongoing pandemic using RSVs of key terms related to the disease can help to inform about the public attention. Results suggest that the data can also be used to assess the response to messages disseminated. The example of the protective masks showed how initial public controversies and contradictions were reflected in the search behaviour of the population, which only informed themselves about protective masks when the curve already flattened.



Lag	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Baden-Württemberg	0,332	0,298	0,282	0,314	0,368	0,422	0,441	0,411	0,348	0,322	0,355	0,391	0,461	0,512	0,523
Bavaria	0,460	0,417	0,394	0,429	0,467	0,506	0,524	0,489	0,430	0,421	0,444	0,474	0,533	0,582	0,564
Berlin	0,475	0,453	0,441	0,491	0,519	0,577	0,548	0,505	0,458	0,394	0,411	0,413	0,459	0,464	0,421
Brandenburg	0,477	0,467	0,392	0,388	0,420	0,452	0,484	0,463	0,412	0,332	0,320	0,327	0,371	0,464	0,511
Bremen	0,602	0,510	0,390	0,406	0,409	0,394	0,395	0,371	0,291	0,243	0,207	0,238	0,257	0,238	0,203
Hamburg	0,258	0,227	0,216	0,233	0,308	0,338	0,374	0,356	0,332	0,354	0,358	0,422	0,471	0,477	0,510
Hesse	0,472	0,423	0,379	0,403	0,472	0,532	0,560	0,550	0,499	0,408	0,403	0,432	0,504	0,565	0,560
Lower Saxony	0,414	0,409	0,378	0,408	0,447	0,487	0,517	0,497	0,439	0,363	0,371	0,390	0,445	0,499	0,506
Mecklenburg Western Pomerania	0,099	0,045	0,061	0,055	0,104	0,113	0,181	0,181	0,150	0,101	0,138	0,186	0,252	0,352	0,400
Northrhine-Westphalia	0,383	0,373	0,374	0,367	0,412	0,457	0,491	0,489	0,446	0,381	0,381	0,420	0,462	0,524	0,562
Rhineland Palatinate	0,356	0,314	0,287	0,329	0,400	0,439	0,446	0,458	0,445	0,401	0,376	0,384	0,403	0,454	0,494
Saarland	0,319	0,273	0,203	0,123	0,140	0,180	0,220	0,247	0,219	0,209	0,161	0,201	0,328	0,491	0,539
Saxony	0,234	0,246	0,243	0,267	0,288	0,308	0,334	0,344	0,342	0,353	0,392	0,405	0,413	0,442	0,485
Saxony-Anhalt	0,310	0,285	0,216	0,271	0,417	0,409	0,379	0,366	0,294	0,176	0,183	0,257	0,311	0,363	0,390
Schleswig Holstein	0,325	0,287	0,250	0,275	0,329	0,356	0,365	0,351	0,442	0,331	0,300	0,319	0,415	0,471	0,525
Thuringia	0,590	0,487	0,392	0,462	0,521	0,553	0,553	0,519	0,419	0,356	0,321	0,395	0,444	0,455	0,412
Germany	0,394	0,327	0,300	0,326	0,383	0,451	0,474	0,466	0,402	0,355	0,367	0,406	0,459	0,510	0,518

.90 to 1.00	Very high correlation
.70 to .90	High correlation
.50 to .70	Moderate correlation
.30 to .50	Low correlation
.00 to .30	negligible correlation

**Figure 6.** Cross-correlation coefficients displaying the linear association between the RSV of the term 'protective mask' and officially reported COVID-19 cases.

15 February	- Beginning of contact tracing of travellers
27 February	- The crisis unit announces to procure more PPE such as protective masks
02 March	- RKI classifies risk as "moderate"
03 March	- Prohibition of PPE exporting
08 March	- First German Covid-19 death
10 March	- Recommendation to cancel events with >1000 participants
11 March	- WHO speaks of a pandemic for the first time
17 March	- RKI classifies risk as "high"
18 March	- Closure of schools and day care centres
18 March	- Chancellor's speaks to the nation on national TV
22 March	- Introduction of contact restriction policy
30 March	- Mask obligation introduced in Austria
31 March	- German government speaks out against mask obligation
02 April	- RKI recommends wearing a mask for everyone
06 April	- Mask obligation introduced in Jena
15 April	- Contact restriction policies extended until May 3 <sup>rd</sup>
17 April	- MoH on national TV: first wave of infection well overcome; Germany should gradually return to normality while staying alert
19 April	- Introduction compulsory wearing of masks in MV and SN
22 April	- Announcement: Mask obligation in all federal states
27 April	- Mask obligation comes into force nationwide

**Figure 7.** Timeline of communication milestones from 15 February to 27 April 2020.

## Data availability

The Google Trends dataset generated for the current study is available and can be replicated at: <https://trends.google.com/trends/>. Daily incidence data on officially reported COVID-19 cases are publicly available: [https://ngeo-corona-npgeo-de.hub.arcgis.com/datasets/ef4b445a53c1406892257fe63129a8ea\\_0?geometry=-19.624%2C46.270%2C35.879%2C55.886](https://ngeo-corona-npgeo-de.hub.arcgis.com/datasets/ef4b445a53c1406892257fe63129a8ea_0?geometry=-19.624%2C46.270%2C35.879%2C55.886).

Received: 30 September 2020; Accepted: 7 March 2021

Published online: 19 March 2021

## References

- European Centre for Disease Prevention and Control (ECDC). *COVID-19 Situation Update Worldwide, as of 3 August 2020* (2020). <https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases>. Accessed 4 August.
- Robert Koch Institute. *COVID-19: Fallzahlen in Deutschland und Weltweit—Stand: 3.8.2020* (2020). [https://www.rki.de/DE/Content/InfAZ/N/Neuartiges\\_Coronavirus/Fallzahlen.html](https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Fallzahlen.html). Accessed 4 August.
- Ferguson, N. Capturing human behaviour. *Nature* **446**, 733–733 (2007).
- Fenichel, E. P. *et al.* Adaptive human behavior in epidemiological models. *Proc. Natl. Acad. Sci.* **108**, 6306–6311 (2011).
- Matrajt, L. & Leung, T. Evaluating the effectiveness of social distancing interventions to delay or flatten the epidemic curve of coronavirus disease. *Emerg. Infect. Dis.* **26**, 1740–1748. <https://doi.org/10.3201/eid2608.201093> (2020).
- Robert Koch Institute (2020). Leitfaden für den Öffentlichen Gesundheitsdienst zum Vorgehen bei Häufungen von COVID-19. (Berlin, 2020).
- Hsu, Y. C., Chen, Y. L., Wei, H. N., Yang, Y. W. & Chen, Y. H. Risk and outbreak communication: Lessons from Taiwan's experiences in the Post-SARS era. *Health Security* **15**, 165–169. <https://doi.org/10.1089/hs.2016.0111> (2017).
- Glik, D. C. Risk communication for public health emergencies. *Annu. Rev. Public Health* **28**, 33–54. <https://doi.org/10.1146/annurev.publhealth.28.021406.144123> (2007).
- McDonnell, W. M., Nelson, D. S. & Schunk, J. E. Should we fear “flu fear” itself? Effects of H1N1 influenza fear on ED use. *Am. J. Emerg. Med.* **30**, 275–282. <https://doi.org/10.1016/j.ajem.2010.11.027> (2012).
- Covello, V. T. Risk communication: An emerging area of health communication research. In *Communication Yearbook 15* (ed. Deetz, S. A.) 359–373 (Sage, 1992).
- Blendon, R. J. *et al.* Public response to community mitigation measures for pandemic influenza. *Emerg. Infect. Dis.* **14**, 778–786. <https://doi.org/10.3201/eid1405.071437> (2008).
- Holmes, B. J. Communicating about emerging infectious disease: The importance of research. *Health Risk Soc.* **10**, 349–360. <https://doi.org/10.1080/13698570802166431> (2008).
- Reynolds, B. & Seeger, M. W. Crisis and emergency risk communication as an integrative model. *J. Health Commun.* **10**, 43–55. <https://doi.org/10.1080/10810730590904571> (2005).
- Lin, L., Savoia, E., Agboola, F. & Viswanath, K. What have we learned about communication inequalities during the H1N1 pandemic: A systematic review of the literature. *BMC Public Health* **14**, 484–484. <https://doi.org/10.1186/1471-2458-14-484> (2014).
- Aven, T. & Renn, O. *Risk Management and Governance: Concepts, Guidelines and Applications* 159–179 (Springer, 2010).
- Lindell, M. K. & Perry, R. W. The protective action decision model: Theoretical modifications and additional evidence. *Risk Anal.* **32**, 616–632. <https://doi.org/10.1111/j.1539-6924.2011.01647.x> (2012).
- Ripberger, J. T. Capturing curiosity: Using internet search trends to measure public attentiveness. *Policy Stud. J.* **39**, 239–259. <https://doi.org/10.1111/j.1541-0072.2011.00406.x> (2011).
- Baker, L., Wagner, T. H., Singer, S. & Bundorf, M. K. Use of the Internet and e-mail for health care information: Results from a national survey. *JAMA* **289**, 2400–2406. <https://doi.org/10.1001/jama.289.18.2400> (2003).
- Carneiro, H. A. & Mylonakis, E. Google trends: A web-based tool for real-time surveillance of disease outbreaks. *Clin. Infect. Dis.* **49**, 1557–1564. <https://doi.org/10.1086/630200> (2009).
- Milnovich, G. J. *et al.* Using internet search queries for infectious disease surveillance: Screening diseases for suitability. *BMC Infect. Dis.* **14**, 690. <https://doi.org/10.1186/s12879-014-0690-1> (2014).
- Strauss, R. *et al.* Investigating the utility of google trends for Zika and Chikungunya surveillance in Venezuela. *BMC Public Health* **20**, 947. <https://doi.org/10.1186/s12889-020-09059-9> (2020).
- Kurian, S. J. *et al.* Correlations between COVID-19 cases and google trends data in the United States: A state-by-state analysis. *Mayo Clin. Proc.* **95**, 2370–2381. <https://doi.org/10.1016/j.mayocp.2020.08.022> (2020).
- Mavragani, A. & Gkillas, K. COVID-19 predictability in the United States using google trends time series. *Sci. Rep.* **10**, 20693. <https://doi.org/10.1038/s41598-020-77275-9> (2020).
- Carr, L. J. & Dunsiger, S. I. Search query data to monitor interest in behavior change: Application for public health. *PLoS ONE* **7**, e48158–e48158. <https://doi.org/10.1371/journal.pone.0048158> (2012).
- Ayers, J. W. *et al.* A novel evaluation of world no tobacco day in latin America. *J. Med. Internet Res.* **14**, e77. <https://doi.org/10.2196/jmir.2148> (2012).
- Ayers, J. W., Althouse, B. M., Noar, S. M. & Cohen, J. E. Do celebrity cancer diagnoses promote primary cancer prevention? *Prev. Med.* **58**, 81–84. <https://doi.org/10.1016/j.ypmed.2013.11.007> (2014).
- Rosenkrantz, A. B. & Prabhu, V. Public interest in imaging-based cancer screening examinations in the United States: Analysis using a web-based search tool. *Am. J. Roentgenol.* **206**, 113–118. <https://doi.org/10.2214/AJR.15.14840> (2015).
- Bentley, R. A. & Ormerod, P. A rapid method for assessing social versus independent interest in health issues: A case study of ‘bird flu’ and ‘swine flu’. *Soc. Sci. Med.* **71**, 482–485. <https://doi.org/10.1016/j.socscimed.2010.03.042> (2010).
- Kostkova, P., Fowler, D., Wiseman, S. & Weinberg, J. R. Major infection events over 5 years: How is media coverage influencing online information needs of health care professionals and the public? *J. Med. Internet Res.* **15**, e107. <https://doi.org/10.2196/jmir.2146> (2013).
- Polkowska, A. *et al.* Increased incidence of Mycoplasma pneumoniae infection in Finland, 2010–2011. *Euro Surveill.* <https://doi.org/10.2807/ese.17.05.20072-en> (2012).
- Muselli, M., Cofini, V., Desideri, G. & Necozone, S. Coronavirus (Covid-19) pandemic: How may communication strategies influence our behaviours? *Int. J. Disaster Risk Reduct.* **53**, 101982. <https://doi.org/10.1016/j.ijdrr.2020.101982> (2021).
- Aksoy, C., Ganslmeier, M. & Poutvaara, P. Public attention and policy responses to COVID-19 pandemic. *MedRxiv* **12**, 348. <https://doi.org/10.1101/2020.06.30.20143420> (2020).
- Husnayain, A., Fuad, A. & Su, E. C.-Y. Applications of google search trends for risk communication in infectious disease management: A case study of the COVID-19 outbreak in Taiwan. *Int. J. Infect. Dis.* **95**, 221–223. <https://doi.org/10.1016/j.ijid.2020.03.021> (2020).
- Hill, S. *et al.* Natural supplements for H1N1 influenza: Retrospective observational infodemiology study of information and search activity on the Internet. *J. Med. Internet Res.* **13**, e36. <https://doi.org/10.2196/jmir.1722> (2011).

35. Rovetta, A. & Castaldo, L. The impact of COVID-19 on Italian web users: A quantitative analysis of regional hygiene interest and emotional response. *Cureus* **12**, e10719. <https://doi.org/10.7759/cureus.10719> (2020).
36. Hinkle, D., Wiersma, W. & Jurs, S. *Applied Statistics for the Behavioral Sciences* (Houghton Mifflin, 2003).
37. Gesis Panel Team. (GESIS Datenarchiv, Köln. ZA5667 Datenfile Version 1.1.0, 2020) <https://doi.org/10.4232/1.13520>.
38. Effenberger, M. *et al.* Association of the COVID-19 pandemic with internet search volumes: A google trends TM analysis. *Int. J. Infect. Dis.* **95**, 192–197. <https://doi.org/10.1016/j.ijid.2020.04.033> (2020).
39. Milinovich, G. J., Williams, G. M., Clements, A. C. A. & Hu, W. Internet-based surveillance systems for monitoring emerging infectious diseases. *Lancet Infect. Dis.* **14**, 160–168. [https://doi.org/10.1016/S1473-3099\(13\)70244-5](https://doi.org/10.1016/S1473-3099(13)70244-5) (2014).

### Author contributions

R.S. developed the concept and design of the study and made substantial contributions to the analysis, interpretation of data and critically revised the manuscript. K.K. performed the analyses, interpreted the data, and wrote the first draft of the article and made substantial contributions the conception of the work. E.L. made substantial contributions to the analysis, interpretation of data and critically revised the manuscript. J.M. revised the last version of the manuscript before submission. All authors have read and approved the manuscript in its current form.

### Funding

Open Access funding enabled and organized by Projekt DEAL.

### Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to R.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021